Heavy-tailed Linear Bandit with Huber Regression (Supplementary Material)

Minhyun Kang[1] **Gi-Soo Kim**[1,2]

[1] Artificial Intelligence Graduate School, UNIST, Ulsan, Republic of Korea
[2] Department of Industrial Engineering, UNIST, Ulsan, Republic of Korea

We present the detailed proof of the result here. For some lemmas we follow the lines of Bastani and Bayati [2020] which prove an analogous bound for the Lasso estimator. For some calculation difference, we present them as well. We indicate it in the corresponding lemmas.

Proof of Lemma 1. Using $\mathbb{E}\left[XX^T\mathbb{1}_{(X\notin U)}\right]$ is semi-positive definite,

$$
\begin{aligned}
\mathbb{E}[XX^T|X\in U] &= \mathbb{E}\left[XX^T\mathbb{1}_{(X\in U)}\right]\cdot\frac{1}{\mathbb{P}(x\in U)}\\
&\preccurlyeq \mathbb{E}\left[XX^T\mathbb{1}_{(X\in U)}\right]\cdot\frac{1}{p}\\
&\preccurlyeq \mathbb{E}\left[XX^T\mathbb{1}_{(X\in U)}\right]\cdot\frac{1}{p}+\mathbb{E}\left[XX^T\mathbb{1}_{(X\notin U)}\right]\cdot\frac{1}{p}\\
&= \mathbb{E}[XX^T]\cdot\frac{1}{p}.
\end{aligned}
$$

□

The following Lemma A states that the size of the set $T_{i,t}$ is $O(\log T)$.

Lemma A (Lemma EC.8 of Bastani and Bayati [2020]). *When* $t\geq (Kq)^2,\ Kq\geq 4,$

$$\frac{1}{2}q\log t < |T_{i,t}| < 2q\log t.$$

Proof of Lemma A. We follow the lines of Lemma EC.8 of Bastani and Bayati [2020]. Let N_t be the largest integer with $t > 2^{N_t+1}Kq$. Then $t\leq 2^{N_t+2}Kq$ and

$$(N_t+2)q\leq |T_{i,t}|\leq (N_t+3)q.$$

For the lower bound, we have

$$\frac{\log(t/Kq)}{\log 2} < N_t+2.$$

Hence,

$$|T_{i,t}|\geq q\frac{\log(t/Kq)}{\log 2}\geq q\log(t/\sqrt{t})=\frac{1}{2}q\log t.$$

Accepted for the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023).

The second inequality follows from $t > (Kq)^2$. For the upper bound, using $N_t + 1 \leq \frac{\log(t/Kq)}{\log 2}$,

$$
\begin{aligned}
|T_{i,t}| &\leq \left(\frac{\log(t/Kq)}{\log 2} + 2\right) q \\
&= \left(\frac{\log(t/Kq) + \log 4}{\log 2}\right) q \\
&= \left(\frac{\log(4t/Kq)}{\log 2}\right) q \\
&\leq 2q \log t.
\end{aligned}
$$

The last inequality follows from $Kq \geq 4$. □

Proof of Lemma 4. We follow the lines of Proposition 2 of Bastani and Bayati [2020]. By the Theorem 2, we have

$$\mathbb{P}\left(\lambda_{\min}\left(\hat{\Sigma}(T_{i,t})\right) \leq \frac{\gamma p}{2}\right) \leq d \exp\left(\frac{-|T_{i,t}|\gamma p}{8}\right).$$

The size of the set $T_{i,t}$ is bounded by

$$|T_{i,t}| \geq \frac{1}{2} q \log t \geq \frac{8}{\gamma p} \log\left(\frac{t^2 d}{\alpha}\right),$$

provided that $q \geq \frac{48}{\gamma t}$ and $t \geq \frac{d}{\alpha}$. Hence, with probability at least $1 - \frac{\alpha}{t^2}$,

$$\lambda_{\min}\left(\hat{\Sigma}(T_{i,t})\right) \geq \frac{\gamma p}{2}. \tag{1}$$

When $q \geq \frac{192}{\gamma p} d^{1/2}$ and $t > \frac{2d+1}{\alpha}$, $|T_{i,t}| \geq 32\lambda_{\min}^{-1}\left(\hat{\Sigma}(T_{i,t})\right) d^{1/2} \log(t^2(2d+1)/\alpha)$. Then, Theorem 1 can be directly applicable with $\tau = \tau_0(|T_{i,t}|/\log(t^2(2d+1)/\alpha))^{1/(1+\delta)}$, $\tau_0 \geq \nu_\delta$. Hence,

$$\mathbb{P}\left(||\hat{\beta}(T_{i,t}) - \beta_i||_2 \leq \left(\frac{\log(t^2(2d+1)/\alpha)}{|T_{i,t}|}\right)^{\delta/(1+\delta)} \cdot 4\lambda_{\min}^{-1}\left(\hat{\Sigma}(T_{i,t})\right) \tau_0 d^{1/2}\right) \geq 1 - \frac{\alpha}{t^2}.$$

Together with (1), when $q \geq 6\left(\frac{32\tau_0 d^{1/2}}{h\gamma p}\right)^{(1+\delta)/\delta}$ and $t \geq \frac{2d+1}{\alpha}$, with probability at least $1 - \frac{2\alpha}{t^2}$,

$$||\hat{\beta}(T_{i,t}) - \beta_i||_2 \leq \frac{h}{4}.$$

□

Proof of Lemma 7. We follow the lines of Lemma EC.14 of Bastani and Bayati [2020]. We have

$$\mathbb{1}_{(r \in \mathcal{A}_{i,t})} = \mathbb{1}_{(A_{r-1})} \cdot \mathbb{1}_{(x_r \in U_i)} \cdot \mathbb{1}_{(r \notin \cup_{i \in [k]} T_{i,t})}.$$

For $n = 0,\ 1,\ 2,\ \ldots$,

$$r \in [(2^n - 1)Kq + 1, 2^n Kq]$$

are forced-sampling time steps and

$$r \in \left[2^n Kq + 1, (2^{n+1} - 1)Kq\right]$$

are not. Let N_t be the largest integer such that $t > 2^{N_t+1}Kq$ as before. Define the intervals

$$V_{1,t} = \left[2^{N_t}Kq + 1, (2^{N_t+1} - 1)Kq\right],\ V_{2,t} = \left[2^{N_t+1}Kq + 1, t \wedge (2^{N_t+2} - 1)Kq\right],$$

and the sum of random variables

$$
\begin{aligned}
M_{i,t} :&= \sum_{r \in V_{1,t}} \mathbb{1}_{(r \in \mathcal{A}_{i,t})} + \sum_{r \in V_{2,t}} \mathbb{1}_{(r \in \mathcal{A}_{i,t})} \\
&< \sum_{r=1}^{t} \mathbb{1}_{(r \in \mathcal{A}_{i,t})} \\
&= |\mathcal{A}_{i,t}|.
\end{aligned}
$$

Both intervals $V_{1,t}$ and $V_{2,t}$ are not containing the forced-sampling time steps and hence we do not update the forced-sample estimator within the intervals. Therefore, we can write

$$\begin{aligned}M_{i,t} &= \sum_{r\in V_{1,t}} \mathbb{1}_{(A_{2^{N_t}}Kq)} \cdot \mathbb{1}_{(x_r\in U_i)} + \sum_{r\in V_{2,t}} \mathbb{1}_{(A_{2^{N_t+1}}Kq)} \cdot \mathbb{1}_{(x_r\in U_i)} \\ &\geq \mathbb{1}_{(A_{2^{N_t}}Kq)} \cdot \mathbb{1}_{(A_{2^{N_t+1}}Kq)} \cdot \sum_{r\in V_{1,t}\cup V_{2,t}} \mathbb{1}_{(x_r\in U_i)}.\end{aligned}$$

The lower bound of cardinality of two disjoint intervals is

$$\begin{aligned}|V_{1,t}\cup V_{2,t}| &= \left(t\wedge 2^{N_t+2}-1\right)Kq - 2^{N_t+1}Kq + \left(2^{N+t+1}Kq - Kq - 2^{N_t}Kq\right) \\ &= \left(t-2^{N_t}Kq - Kq\right)\wedge\left(3\cdot 2^{N_t}Kq - 2Kq\right) \\ &> \left(\frac{t}{2}-Kq\right)\wedge\left(\frac{3}{4}t-2Kq\right) \\ &> \left(\frac{t}{2}-\frac{t}{80}\right)\wedge\left(\frac{3}{4}t-\frac{t}{40}\right) \\ &= \frac{39}{80}t.\end{aligned}$$

The first inequality follows from $t\leq 2^{N_t+2}Kq$. The last inequality follows from $t>(Kq)^2$ and $q>80$. The upper bound of the cardinality of two disjoint intervals is

$$\begin{aligned}|V_{1,t}\cup V_{2,t}| &< t-2^{N_t}Kq-Kq \\ &< t-\frac{t}{4}-Kq \\ &< \frac{3}{4}t.\end{aligned}$$

The probability of two events is bounded by

$$\begin{aligned}\mathbb{P}\left(A_{2^{N_t}Kq} \text{ and } A_{2^{N_t+1}Kq}\right) &\geq 1-\frac{2K\alpha}{(t/4)^2}-\frac{2K\alpha}{(t/2)^2} \\ &= 1-\frac{32K\alpha}{t^2} \\ &> 1-0.01.\end{aligned}$$

The last inequality is from $t^2>(Kq)^4$ and $\alpha\in(0,1)$. Hence, we have

$$\begin{aligned}\mathbb{E}[M_{i,t}] &\geq \mathbb{P}\left(A_{2^{N_t}Kq} \text{ and } A_{2^{N_t+1}Kq}\right)p|V_{1,t}\cup V_{2,t}| \\ &\geq 0.48tp.\end{aligned}$$

The Hoeffding's inequality implies,

$$\begin{aligned}\mathbb{P}\left(\mathbb{E}[M_{i,t}]-M_{i,t}\geq\eta^2\right) &\leq \exp\left(-\frac{2\eta}{|V_{1,t}\cup V_{2,t}|}\right) \\ &\leq \exp\left(-\frac{8\eta^2}{3t}\right).\end{aligned}$$

Let $\eta=0.23tp$. Then

$$\begin{aligned}\mathbb{P}(M_{i,t}<0.48tp-0.23tp) &\leq \exp\left(-\frac{8}{3}t(0.23p)^2\right) \\ &\leq \exp(-tp^2/9).\end{aligned}$$

Since $M_{i,t}\leq|\mathcal{A}_{i,t}|$,

$$\mathbb{P}\left(|\mathcal{A}_{i,t}|<\frac{tp}{4}\right)\leq\exp(-tp^2/9)\leq\frac{\alpha}{t^2},$$

provided that $t\geq\frac{1}{\alpha}$ and $q\geq\frac{54}{p}$. $\square$

We now provide the proof of the expected regret bound.

Proof of Theorem 4. Lemma EC.19 of Bastani and Bayati [2020] states that the upper bound of expected regret can be decomposed into

$$\sum_{t=1}^{T} \mathbb{E}[r_t] = \sum_{t}^{T} \mathbb{E}[x^T \beta_{a^*(t)} - x^T \beta_{a(t)}]$$
$$\leq 2 \sum_{i \in \mathcal{D}} \mathbb{P}(||\hat{\beta}(S_{a^*(t),t-1}) - \beta_{a^*(t)}||_2 > \Delta) + 2 \sum_{i \in \mathcal{D}} \mathbb{P}(||\hat{\beta}(S_{a(t),t-1}) - \beta_{a(t)}||_2 > \Delta) + 4\Delta^2 K C_0$$

for $\Delta > 0$. From Lemma 7 with $\alpha = (2d+1)t$, we have

$$\mathbb{P}\left(||\hat{\beta}(S_{i,t}) - \beta_i||_2 \geq \left(\frac{4}{pt}\log t\right)^{\delta/(1+\delta)} \frac{32\tau_0 d^{1/2}}{\gamma p} \right) \leq \frac{3(2d+2)}{t}$$

for $i \in K_{opt}$. Let $\Delta = \left(\frac{4}{pt}\log t\right)^{\delta/(1+\delta)} \frac{32\tau_0 d^{1/2}}{\gamma p}$ then,

$$\mathbb{E}[r_t] \leq \frac{12K(2d+1)}{t} + 4\left(\frac{32\tau_0}{\gamma p}\right)^2 d \left(\frac{4}{pt}\log T\right)^{\frac{2\delta}{1+\delta}} K C_0.$$

The cumulative regret is bounded by

$$\sum_{t=1}^{T} \mathbb{E}[r_t] \leq 12K(2d+1)(\log T + 1) + 4^7 d \left(\frac{\tau_0}{\gamma}\right)^2 \frac{1}{p^3} K C_0((\log T)^2 + \log T)$$

when $\delta = 1$ and

$$\sum_{t=1}^{T} \mathbb{E}[r_t] \leq 12K(2d+1)(\log T + 1) + 64^2 16^{\frac{\delta}{1+\delta}} d \left(\frac{\tau_0}{\gamma}\right)^2 \frac{1}{p^{\frac{2+4\delta}{1+\delta}}} K C_0 \left(\frac{1+\delta}{1-\delta}\right) T^{\frac{1-\delta}{1+\delta}} (\log T)^{\frac{2\delta}{1+\delta}}$$

when $0 < \delta < 1$.

□

References

Hamsa Bastani and Mohsen Bayati. Online decision making with high-dimensional covariates. *Operations Research*, 68(1): 276–294, 2020.